

映像によるライフログ

中村 裕一*

映像には多くの情報を記録することができ、ライフログのデータとして優れている反面、閲覧に時間がかかり、一覧性も悪いなどの欠点もある。そのため、メディア技術を用いて、ライフログ映像の加工を行うことが必須となる。本稿では、記憶補助から集団活動の解析や補助まで、様々な目的に対して、種々の画像・映像処理を行って、ライフログ映像の活用を支援する手法について紹介する。

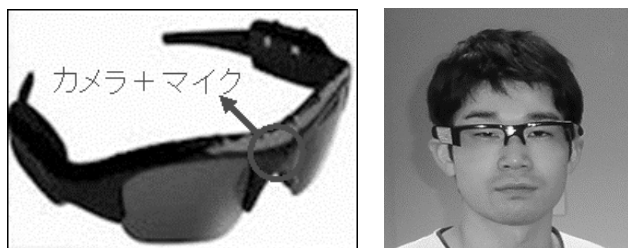
キーワード：ライフログ、映像、主観映像、グループのログ、シーンの再構成

1. 映像によるライフログの取得

個人の体験をライフログとして記録する場合、ビデオカメラなどによる映像記録が最も強力な方法の一つとなる。何か重要なことを意図的に記録する場合には、スチルカメラのようなデバイスが向いているが、そうでない場合には、映像を常時撮り流すような記録方法が適している。後になってどの場面が重要になるか、あらかじめわからないからである。

近年、映像の撮影・記録デバイスが小型化したことにより、撮影の負担が軽減されている。実際に、頭部や肩、胸にビデオカメラやスチルカメラを装着し、装着者の前や周辺を長時間記録する研究や実践が行われてきている。図1はカメラが内蔵された眼鏡の例である。このような方法で得られる映像は、装着者の視点から見たシーンを記録していることから、個人視点映像（Point of View）や主観映像（First Person Vision）等と呼ばれる。なお、SenseCam¹⁾のように、胸の前にぶら下げたスチルカメラを用いて定期的に、例えば、30秒や1分毎に静止画像を記録していく方法も用いられている。この方法では、記録できる情報量は映像のように多くないが、記録される側の理解が得やすいこと、データ量が小さくなることが利点となる。

その他にも、部屋やオフィスなどの環境に固定されたカ



(a) サングラス型カメラ (b) 装着時の様子

図1 カメラの装着

*なかむら ゆういち 京都大学 学術情報メディアセンター
〒606-8501 京都市左京区吉田本町
Tel. 075-753-7477 (原稿受領 2012.12.13)



図2 俯瞰映像と個人視点映像の併用（左部分の線は行動軌跡）

メラを併用することも考えられる。例えば、全方位ミラーや魚眼レンズを用いて周囲を広く撮影すれば、行動者本人からは直接見えない部分、例えば背後の様子も記録できる。それを図2のように個人視点の映像記録と対応づければ、場の状況を理解しやすいデータとなる²⁾。

また、映像と同時に種々のセンサを併用して、その場の状況に関する情報（位置、温度、その他）や本人の生理的状态（例えば、発汗、脈、筋電）等を記録することも良く行われる。見えない、つまり映像として記録できない情報を補うことによって、各々の場面の重要性を知る手がかりにしたり、検索を簡単にできる。

2. ライフログ映像の利用

2.1 ライフログの目的

①個人の記憶補助

ライフログ映像の最もわかりやすい用途は、個人の記憶補助だろう。これを、出来事が起ったことを記憶するエピソード記憶と、それぞれの出来事における重要な情報を記憶する意味記憶に大まかに分けて考えてみよう。認知症を患う人や高齢者にはエピソード記憶の補助が重要なテーマとなる。健常者にとっては、意味記憶としてのライフログ

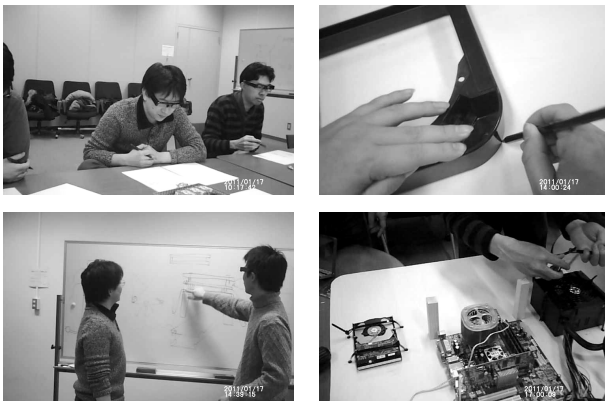


図3 集団行動におけるライフログ映像

の活用が大いに考えられる。

前者の例としては、記憶に障害を抱えた人に SenseCam を 1 年弱の間利用してもらったケーススタディがある³⁾。毎日記録されるデータを介護者といっしょに閲覧することによって、出来事を覚える度合い、また、時間が経ったときに想起できる度合いの両方に良い効果があったと報告されている。健常者にとっても関わりの深いものとしては、なくしもの検索が考えられるだろう。ライフログ映像があれば、何かを置き忘れた場面を検索すれば良いことになる⁴⁾。

その他にも、熟練やコツを必要とする高度な技を記録することや、作業などを記録することによって行動や事故の解析を行うことなど、様々な利用が考えられる。こういった用途のために映像を撮影すること自体は従来から良く行われてきたが、ライフログ的な映像取得とその利用の検討はまだ十分に行われていない。

②グループの体験記録

複数の人間によってライフログ映像を取得できれば、個人のライフログにはなかった活用範囲が広がる。例えば、体験型の学習や見学、ワークショップ（例を図3に示す）などを考えるとわかりやすい⁵⁾。これらの活動では、コストをかけて準備された環境で、貴重な体験や発見を発生する。これらを記録し、統合することで、まず、場の「総合的な記録」が得られる。つまり、各人が異なる視点や立場から対象を見ているため、それらを集めることにより、より総合的、俯瞰的に活動の場を記録することができる。また、「体験を比較する」ことができる。つまり、異なる立場や視点の人から得た記録を比較することによって、貴重な情報が得られる。例えば、参加者にとっては、他の参加者の記録によって自分の体験の補完を行ったり、自分や他人のやり方の違いを知る手がかりとなる。また、参加者間のコミュニケーションを解析するための良いデータともなる。

③その他のライフログ

図4のように、お互いに自分の撮っている映像を相手に送り、それを見ながら対話をする映像対話型のコミュニケーションがある⁶⁾。実際に、ヘルメットに付けたカメラの映像を携帯電話の回線を用いて送るシステムなどが実用

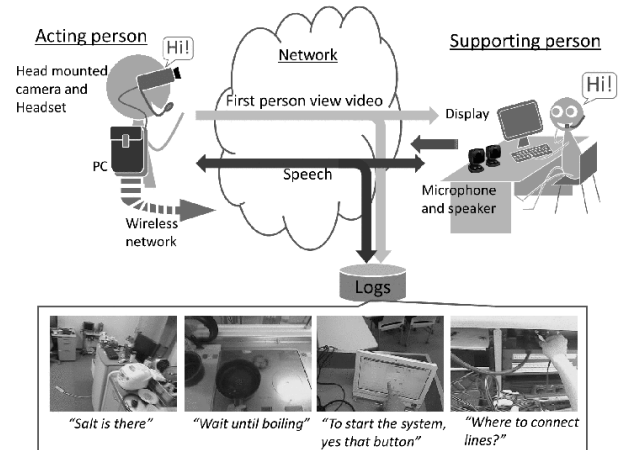


図4 映像対話型行動記録

化されている。このような手法は、直接現地に呼ぶことが難しい専門家に遠隔地から指示を出してもらおう場合や、危険な場所において、外にいる人と情報を共有しながら作業を行う場合等に活用できる。このような映像にもライフログとしての特徴や価値がある。当事者が作業などの状況を想起するために用いたり、事故などの原因を探ったり、第三者に作業を説明する資料にすることなどが期待できる。

人間だけではなく動物の行動解析にも映像によるライフログ技術が用いられ始めている。海洋生物、野生動物、家畜などにセンサを装着してその生態を記録するバイオロギングの分野があるが⁷⁾⁸⁾、それに映像を用いるものである。図5は絶滅危惧種となっているアオウミガメの生態解明や保護を目的として、石垣島で取得された映像の一コマであり、甲羅に GPS 等のセンサとともにビデオカメラを装着することによって撮影されている。このような映像により、どのような食餌行動をとっているか、どのように睡眠をとっているかなど、貴重な情報が得られている。



図5 バイオロギングにおける映像取得（アオウミガメ）

2.2 ライフログ映像による支援

これまであげてきたライフログによる支援の内容とその要件について整理すると以下のようなになる。

思い出すための閲覧：エピソード記憶をライフログによって補完する。長時間の映像データを閲覧するための要約や、重要な部分の効率的な検索が必要となる。

記憶の整形、再体験のための閲覧：エピソード記憶の整形をするための補助を行う。ライフログデータを閲覧しながら正確なエピソード記憶を作り直すことは、記憶に障害のある人だけでなく、健常者にとっても助けとなる場合がある。また、他人に自分の体験を説明するためにそれをを用いることができる。

新しい気づきや発見のための閲覧：記録時には気づかなかった情報を得る。思い出すための閲覧と同等の機能が必要となる他、記録のわかりやすい提示や種々の観点からの加工、データマイニングなどを含む統計的処理などが有用となる。

他人の体験との総合・比較・補完：複数人のライフログ映像をまとめて提示したり、加工したり統合して提示することによって、集団としての体験をまとめたり、個人間の差異を明らかにすることができる。

3. ライフログ映像に対する処理

3.1 重要なシーンの検出

エピソード記憶、また、それに付随する意味記憶を補助することを考えると、位置・場所、行動をライフログ映像から検出する必要がある。従来、ニュース映像やドラマ映像に対するインデキシングや要約手法が提案されてきており、ライフログ映像の処理にも多くの手法が適用できる。ただし、ライフログ映像の特殊な点として、(1)個人視点から撮り流された動きの激しい(カメラワークを含んだ)映像であり、ショットなどの切目が無い、(2)付加的なセンサ情報(例えば、位置情報や装着者の生体情報)を利用できる場合がある等の違いがある。特に、(1)に関して行動の検出手法が異なってくる。

以下では、「場所・移動」として、カメラの移動を、重要な「行動」として「手の動き」と「見る行動」を検出する手法を紹介する。「見る行動」は記録をとった時点での本人の興味や意図を表わす良いインデックスとなる。興味を持って意図的に見たものは記憶に残りやすい上に、視点の停留や見回しが起った場面では、その時の状況を再認識しやすいからである。

また、本稿ではふれないが、生理的な内部状態、例えば、興奮や緊張なども良いインデックスとなることが期待できる。脈、呼吸、筋電位、脳波、脳血流等のウェアラブルなセンサも種々開発されており、これらを利用できる。

①カメラの自己運動

最も基本的なものは本人(カメラ装着者)の運動(移動)となる。カメラの動きが映像の幾何的变化として現れるため、それを検出すれば良い。

その方法の一つは、現在のフレームと直前・直後のフレームの輝度値の相関を用いる方法である。具体的には、画像を各画素の輝度値を要素とする多次元のベクトルと見なした場合の、ベクトル間の距離またはそれを正規化したもの



図6 注視・注目の検出例

を小さくする。式1に示したように、映像中の時刻 t の画像を基準 I_t とし、 Δt フレーム後の画像 $I_{t+\Delta t}$ との間で $Err(I_t; \mathbf{A})$ を最小にするような幾何変換 \mathbf{A} を求める。

$$Err(I_t; \mathbf{A}) = \sum_{\mathbf{x}} (I_{t+\Delta t}(\mathbf{x}) - I_t(\mathbf{x}; \mathbf{A})) \quad (1)$$

\mathbf{A} の幾何変換モデルとしては、2次元アフィン変換(線形変換と平行移動)や、3次元剛体モデルに対するカメラ移動を仮定するモデルなどが提案されている^{註1)}。最近では、SIFT等の局所特徴の対応関係を用いてより詳細に対応関係を求めることも行われる。

②移動・見直し

得られたカメラの自己運動から視点の停留や見直しなどが検出できる。例えば、歩行による前進移動時には、幾何変換のパラメータが細かな上下の移動をともなったスケールの拡大を示す。同様に、方向転換や横を向きながらの移動があれば、パラメータがそれに応じた特徴的な振る舞いを示す。

③注目箇所

背景の動きと異なる動きが中央付近に現れば、それが注目対象の候補となる¹²⁾。そのためには、 $|I_{t+\Delta t}(\mathbf{x}) - I_t(\mathbf{x}; \mathbf{A})|$ が大きくなる部分(\mathbf{x})が視界の中央付近にとどまり続けることを検出すれば良い。また、視点の停留が起っている場合もその候補となる。実際に検出された注目シーンと注目対象(矩形で囲われた部分)を図6に示す。

④手を使う動作: 物を取る・置く

手領域の検出と追跡は、色情報やオブティカルフロー^{註2)}を用いることによりある程度行うことができる。興味深い現象として、物を取ったり、置いたりする際には特徴的な頭部の動きが現れる。そのため、頭部の動き(カメラの自己運動)と手の動きの相互関係を用いれば、物を取る・置く動作の検出精度が上がるのがわかってきている。

3.2 代表画像による閲覧

空間展開

長い映像の要約・閲覧手法として、ショット分割を行って、そこから代表画像を選びだし、それを空間的に並べることが一般的に行われてきた。ライフログ映像にはショットのような切目が無いが、移動・動作・注目など、重要なシーンを検出することによって切目や代表画像とすることができる。例えば、図7は移動と注目をもとにライフログ映像を要約した例である。

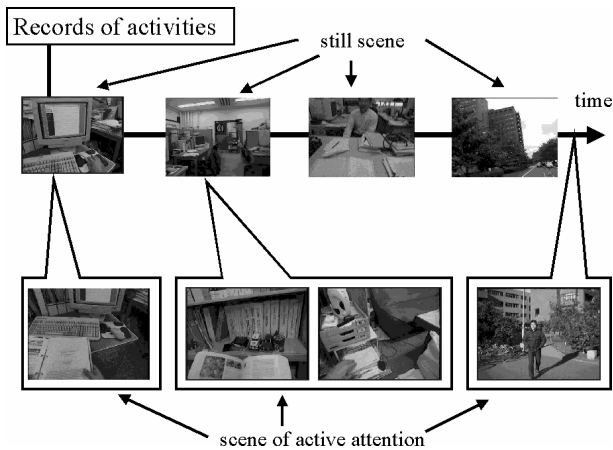


図7 ライフログ映像の空間展開

思考の道具としての利用

図8のようにサムネイル画像を使って、考えをまとめるためのオンライン環境を与えることもできる。筆者らの実験では、参加者によって好む素材が変わること、個人視点映像を好む人もいることがわかってきている⁵⁾。



図8 サムネイル画像やメモを用いた思考整理(一部分を切り出して表示している)

3.3 貼りあわせによる広視野化

頭部や体に装着しても自然な行動を妨げないよう、ライフログ映像用のカメラは小型のものになる。そのため、十

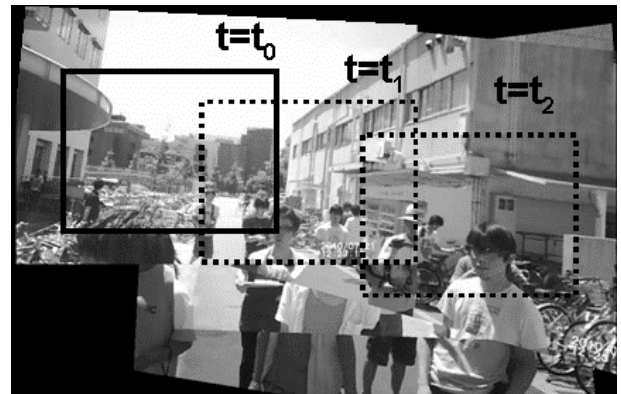


図9 画像の貼り合わせ結果(動いている人や変化している部分はうまく貼り合わせを行うことができない)。t=t₀, t₁, t₂は各々異なる時刻の視野を示す。

分な解像度で周囲の状況を記録することができない。つまり、視野を狭くするか、解像度を犠牲にして広角の映像を得るか、どちらかを選択することになる。一般的には、エピソード記憶を補助する場合には視野が、意味記憶を補助する場合には解像度が重要となると考えられるが、画像処理により擬似的に両方を満たすことができる場合もある。

撮影時のカメラの自己運動または映像の隣接するフレーム間の幾何学的関係がわかっているならば、画像を貼り合わせる(mosaicingまたはstitchingと呼ばれる)ことができる¹⁰⁾¹¹⁾。図9に例を示すが、視野の狭さが緩和され、周囲の状況がわかりやすくなる。貼り合わせの出来映えは、カメラの動き、シーンの3次元構造、移動物体の有無に依存する。原理的には、カメラ中心まわりの回転のみがある場合、シーンが平面(または遠景のみ)の場合には、誤差のない貼り合わせが可能である。

3.4 複数ライフログ映像の閲覧

複数の人間のライフログ映像を取得する際にその相互関係が重要な意味を持つ場合も多い。しかし、複数の動画を並べるだけでは、閲覧するのがつらく、理解しづらいものになる場合が多い。

複数視野の埋め込み

映像間の関連がわかりやすいように閲覧する方法として、シーンを俯瞰する映像に個人視点映像を埋め込む手法があげられる。これにより、誰がどこでどのように何を見たかがわかりやすくなる。図10の例は、天井からの俯瞰映像に、各人物の視点から撮られた映像を各人物の頭部の位置・方向に合わせて埋め込んだものであり、中央の物体(マニュアル)に全員が注目していることや、その見え方等がわかる。このような俯瞰映像は必ずしも上から撮影したものである必要はなく、各々の俯瞰映像におけるおおよその頭部の位置と視線の方向がわかっているならば、映像を埋め込むことができる。

視線に注目した再構成

複数人が同時に活動したり、違う時刻に同じものを見たりする場合に、複数人の視線・視野の関係を解析したいと



図 10 俯瞰映像への個人視点映像の埋め込み（各人の頭の後ろに映像が埋め込まれている）

いう要求がある。例えば、説明者が効率的に聴衆の注意・注目を集めているか、展示物が効果的に注意を惹いているかなどを確認したり、皆の注意を惹いた部分を発見することの需要は大きい。



図 11 視線に注目した再構成

図 11 はある人の視野にその他の人の視線を擬似的に書き込んだ例である。本人以外の視線は円錐で表現されており、これが集まれば集まるほどその点が皆に注目されていることを直観できる¹³⁾。このような合成を行うためには、各人の視点（カメラ位置）とその方向を知る必要があるが、この例では SLAM¹⁴⁾¹⁵⁾と AR マーカ¹⁶⁾を使って、各人のカメラ位置・方向を共通の座標系に変換している。

4. おわりに

本稿では、映像によるライフログの取得と利用について紹介した。ここで主に焦点をあてたのは、短時間で理解できるように整理する支援、複数の映像を総合して閲覧する支援など、人間が直観的にわかりやすいようにライフログ映像を加工する処理である。このような処理によって、ライフログ映像からの情報の検索や発見が簡単になり、より多くの目的で利用されることを期待できる。また、本稿で

はほとんど触れなかったが、ライフログ映像からの一般的な行動や環境の自動認識によっても重要な情報が得られる。それにはまだ多くの課題が残っているが、近年、First Person Vision の処理が盛んに研究されており、今後の進展が大いに期待できる。

註・参考文献

- 註 1) 撮影されている世界が 3 次元的に奥行きを持ち、形状の時間的変化がないと仮定する。それを任意の 3 次元位置から撮影したことを仮定する。一般的には、3 次元計測を含む問題となるため、良い精度を得るのは簡単ではない。そのために様々な方法が提案されている。
- 註 2) 画像中の各微小要素（画素やブロック）などの動きベクトルを求めたものをオプティカルフローと呼ぶ。微小要素のとりかたや計算方法の異なる様々な手法がある。
- 1) S. Hodges, et.al: SenseCam: A Retrospective Memory Aid; UbiComp 2006, Lecture Notes in Computer Science Volume 4206, 2006, pp.177-193.
 - 2) S.Kubota, Y.Nakamura, Y.Ohta: Detecting Scenes of Attention from Personal View Records; Proc. IAPR Workshop on Machine Vision and Applications, 2002. pp.209-213.
 - 3) E. Berry, et.al: The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: a preliminary report, Neuropsychol Rehabil. 2007. vol.17, no.4-5: pp.582-601.
 - 4) 小泉, 中村, 佐藤: 探し物検索のための個人視点映像からの手掛かり発見; 信学技報 MVE2010-41, 2010, pp.105-109.
 - 5) 近藤ほか: 個人視点映像を用いた気づき体験の回想と整理支援 ~フィールド調査における問題発見を通じて~; 信学技報 PRMU2010-128, 2010, pp.13-18.
 - 6) 京都大学フィールド情報学研究会編, フィールド情報学, 2009, 共立出版.
 - 7) 日本バイオロギング研究会編, バイオロギングー最新科学で解明する動物生態学一, 京都通信社, 2009.
 - 8) 小泉ほか: 映像対話型行動記録におけるモダリティ間の関係と凝集性, 信学技報 HCS2012-33, 2012. vol.112, no.176, pp.1-6.
 - 9) 角ほか: 共有体験を通じたメタ認知に対する複数視点映像の効果; 情報処理学会論文誌, 2008. vol.49, no.4, pp.1637-1647.
 - 10) Y.Nakamura, J. Ohde, Y.Ohta: Structuring Personal Experiences - Analyzing Views from a Head-mounted Camera; Proc. Int'l Conference on Multimedia and Exposition 2000, 2000, pp.CD-ROM TP10-5.
 - 11) R. Szeliski: Video mosaics for virtual environments; IEEE Computer Graphics and Applications, March 1996, pp.22-30.
 - 12) M.Brown and D.Lowe: Automatic Panoramic Image Stitching using Invariant Features; International Journal of Computer Vision, 2007. vol.74, no.1.
 - 13) 高瀬, 近藤, 小泉, 中村: 共同注視状況における複数人物頭部カメラの位置姿勢推定, 電子情報通信学会 HCG シンポジウム 2012, 2012, II-1-1, 2012, pp.22-28.
 - 14) H.Durrant-Whyte, T.Bailey: Simultaneous Localization and Mapping(SLAM): Part I The Essential Algorithms; Robotics and Automation Magazine 2006, vol.13, no.2, pp.99-110.
 - 15) G.Klein, D.Murray: Parallel Tracking and Mapping for Small AR Workspaces; The 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007, pp.225-234.
 - 16) H.Kato, M.Billinghurst: Marker tracking and hmd calibration for a video-based augmented reality conferencing system; In Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, 1999.

Special feature: Logs for practical use. Lifelog through Videos. Yuichi NAKAMURA (Academic Center for Computing and Media Studies, Kyoto University, Sakyo, Kyoto,606-8501 JAPAN)

Abstract: Videos are good lifelog media that can capture various kinds of rich information. Videos have,however, certain drawbacks, e.g., reviewing is time consuming, browsing is not easy, etc. Media processing is essential to overcome those problems.This paper introduces media processing techniques for supporting better utilization of lifelog videos for various kinds of purposes such as memory aids,analysis of group activities, etc.

Keywords: lifelog / video / first person vision / group-log / view reconstruction